PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

AJ12

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/27

A2

(11) International Publication Number: WO 90/12874

(43) International Publication Date: 1 November 1990 (01.11.90)

(21) International Application Number: PCT/US90/02144

(22) International Filing Date: 19 April 1990 (19.04.90)

(30) Priority data:
341,990 21 April 1989 (21.04.89) US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 Cambridge Park Drive, Cambridge, MA 02140 (US).

(72) Inventors: SHAW, Gray ; 42 Burlington Road, Bedford, MA 01730 (US). VELDMAN, Geertruida ; 60 Woodmere Drive, Sudbury, MA 01776 (US). WOOTERS, Joseph ; 40 Fairbanks Street, Brighton, MA 02135 (US).

(74) Agent: CSERR, Luann; Genetics Institute, Inc., Legal Affairs Department, 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: CYSTEINE ADDED VARIANTS OF POLYPEPTIDES AND CHEMICAL MODIFICATIONS THEREOF

(57) Abstract

Cysteine added variants ("CAVs") of polypeptides are provided having one or more cysteine residues substituted for selected naturally occurring amino acid residues, or inserted into the polypeptide sequence, and preferably being further modified by deletion of certain N-terminal amino acids. Such CAVs may be additionally modified by the coupling of sulfhydryl reactive compounds to the introduced cysteine residue(s) without loss of bioactivity to produce selected homogeneously modified IL-3, G-CSF or EPO and improved pharmaceutical compositions containing the same.

```
    1                                                  10
ATG GCT CCT ATG ACT CAA ACT ACT TCT TTA AAA ACT TCT
Met Ala Pro Met Thr Gln Thr Thr Ser Leu Lys Thr Ser

                                                       25
TGG GTA AAC TGT TCT AAC ATG ATC GAT GAA ATT ATA ACA
Trp Val Asn Cys Ser Asn Met Ile Asp Glu Ile Ile Thr

CAC TTA AAG CAG CCA CCT TTG CCC TTG CTG GAC TTC AAC
His Leu Lys Gln Pro Pro Leu Pro Leu Leu Asp Phe Asn

    40
AAC CTC AAT GGG GAA GAC CAA GAC ATT CTG ATG GAA AAT
Asn Leu Asn Gly Glu Asp Gln Asp Ile Leu Met Glu Asn

            55
AAC CTT CGA AGG CCA AAC CTG GAG GCA TTC AAC AGG GCT
Asn Leu Arg Arg Pro Asn Leu Glu Ala Phe Asn Arg Ala

                    70
GTC AAG AGT CTG CAA AAT GCA TCA GCA ATT GAG AGC ATT
Val Lys Ser Leu Gln Asn Ala Ser Ala Ile Glu Ser Ile

                                85
CTG AAA AAT CTG CTG CCA TGT CTG CCC CTG GCC ACA GCT
Ile Lys Asn Leu Leu Pro Cys Leu Pro Leu Ala Thr Ala

                                                   100
GCA CCC ACC AGG CAT CCA ATC CAT ATC AAG GAT GGT GAC
Ala Pro Thr Arg His Pro Ile His Ile Lys Asp Gly Asp

                                                       115
TGG AAT GAA TTC CGC CGC AAA CTG ACC TTC TAT CTG AAA
Trp Asn Glu Phe Arg Arg Lys Leu Thr Phe Tyr Leu Lys

ACC CTG GAG AAT GCT CAG GCT CAG CAG ACC ACC CTG AGC
Thr Leu Glu Asn Ala Gln Ala Gln Gln Thr Thr Leu Ser

130
CTC GCG ATC TTC TAG
Leu Ala Ile Phe Stop
```

+ See back of page

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Fasso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo			SN	Senegal
CH	Switzerland	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
DE	Germany, Federal Republic of	LK	Sri Lanka	TG	Togo
DK	Denmark	LU	Luxembourg	US	United States of America

CYSTEINE ADDED VARIANTS OF POLYPEPTIDES AND CHEMICAL MODIFICATIONS THEREOF

TECHNICAL FIELD

This invention relates generally to polypeptides modified by the attachment of compounds having sulfhydryl reactive groups, improved methods for producing such modified polypeptides and improved compositions containing them. The invention relates particularly to three modified polypeptides (IL-3, G-CSF and EPO), to which sulfhydryl reactive compounds, including polymers, may be attached at selected positions in the polypeptide that have been modified by the insertion of cysteine residues or the substitution of cysteine residues for other residues.

BACKGROUND

The desirability of modifying biologically active and therapeutically useful polypeptides with a variety of compounds, such as the hydrophilic polymer polyethylene glycol (PEG), to enhance their pharmacokinetic properties has been noted. See, e.g., the discussion of the art in this area of polypeptide modification in published PCT patent application WO87/00056, in U.S. Pat. No. 4,179,337, which discloses conjugating water soluble polypeptides such as enzymes and insulin to PEG or PPG, and in U.S. Pat. No. 4,766,106, which discloses conjugating ordinarily water insoluble beta-interferon, interleukin-2, or immunotoxins to PEG homopolymers or polyoxyethylated glycerol.

Such modification can reduce adverse immune response to the polypeptide, increase the solubility for use in pharmaceutical preparations and maintain a desirable circulatory level of such polypeptide for therapeutic efficacy.

One problem not addressed by the art in this area involves the extent to which a polypeptide can be modified by attachment of compounds having reactive groups that will covalently bond to certain amino acid residues of the polypeptide. For example, modification of a polypeptide with PEG or similar polymers, can result in random attachment of the polymer at the amino terminus of the polypeptide and/or at one or more lysine residues in the amino acid sequence of the protein. Because more than one PEG group can attach to the polypeptide, the resultant composition may contain a heterogeneous mixture of "PEGylated" polypeptide; some polypeptides having only one PEGylated site, others having more than one PEGylated site. Such heterogeneity in composition is undesirable for pharmaceutical use. Furthermore, the non-specificity with regard to the site(s) of attachment of compounds such as PEG to the polypeptide can result in loss of biological efficacy of the polypeptide stemming from undesirable attachment to a polypeptide site required for biological activity. United States Patent 4,904,584 addresses the foregoing by providing materials and methods for site specific covalent modification of polypeptides by lysine insertion, removal, and/or replacement. However, we have determined that the use of lysine as the attachment site for modification, for example, by PEGylation, may be disadvantageous because not all

modifications may result in biologically active compounds and because steps must be taken to prevent PEGylation at N-termini in cases where N-terminal PEGylation is not desired.

SUMMARY OF THE INVENTION

This invention provides materials and methods for site specific covalent modification of polypeptides, particularly and preferably human IL-3, granulocyte colony stimulating factor (G-CSF) and erythropoietin (EPO) polypeptides, permitting the production of compositions comprising homogeneously cys modified IL-3s, G-CSFs and EPOs and pharmaceutical compositions containing the same. "Homogeneously cys modified" as the term is used herein means substantially consistently modified only at specific, inserted or substituted cysteine residues. A homogeneously modified IL-3 for example, includes an IL-3 composition which is substantially consistently modified at position 6 (using the convention of counting from the N-terminus of the mature protein) by the insertion of cysteine in place of the threonine of natural IL-3, but not at other positions.

Thus, this invention first provides cysteine added variants ("CAVs") of IL-3, G-CSF and EPO. CAVs of this invention encompass IL-3, G-CSF and EPO muteins that contain at least one additional cysteine residue compared to the corresponding naturally occurring or previously known IL-3, G-CSF and EPO. The cysteine residue(s) are introduced into the peptide structure of the CAVs at one or more amino acid positions in the natural or previously known counterpart.

In the case of human IL-3, we have determined that the naturally occurring cysteine residues at positions 16 and 84 form a disulfide bridge, essential to preserving the desired biological activity of the polypeptide. For the addition of novel cysteines, some positions within the polypeptide, such as position 15 and 51 are unsuitable; cysteines introduced at these positions give rise to human IL-3 polypeptides with substantially reduced biological activity. However, certain substitutions or deletions of residues 1-14 do not significantly diminish the desired biological activity of IL-3. Therefore, a preferred region of novel cysteine introduction into the polypeptide is within positions 1-14 inclusive. Currently, positions 6-12 inclusive are especially preferred sites for cysteine introduction. The subsequent attachment of sulfhydryl reactive compounds, including polymers, as discussed below, to the novel cysteines added at selected positions within this region will not result in any significant loss of biological activity.

By "cysteine added variant" as the term is used herein, we mean variants of IL-3, G-CSF and EPO that are modified in amino acid structure relative to naturally occurring or previously known counterparts such that at least one cysteine residue is inserted into the natural or previously known sequence and/or is used to replace a different amino acid within that sequence.

Additionally, with respect to IL-3, the native or "natural" IL-3 sequence, with an added initiator methionine for bacterial expression, may be further modified such that the first alanine is deleted at

the N-terminus of the mature polypeptide, altering the amino terminal sequence from MET*ALA*PRO to MET*PRO (the "mp" mutein). For the "mp" mutein, such N-terminus modification permits more consistent removal of the N-terminal methionine. As is already known, in bacterial expression systems, cleavage at the N-terminal methionine occurs. Likewise, the native EPO N-terminal sequence (with the added MET) begins MET*ALA*PRO and it may prove advantageous to delete the first alanine to obtain an mpEPO mutein. With regard to G-CSF, the natural human N-terminal sequence begins with MET*THR*PRO (with the MET added for bacterial production) and it may be desirable to delete this N-terminal threonine to advantageously obtain a mpG-CSF mutein.

Alternatively, the natural IL-3 sequence may be further modified such that the first two amino acids at the N-terminus of the mature polypeptide are deleted, leaving a terminus beginning with MET*THR*GLN*THR* (the "m3" mutein). For the "m3" mutein, such N-terminus modification permits one to take advantage of the methionine at position 3 in the naturally occurring human IL-3 molecule, as the initiator methionine.

The CAVs of this invention make it possible to produce homogeneous, biologically active IL-3, G-CSF and EPO compositions substantially specifically and consistently modified at selected positions with sulfhydryl reactive compounds (described hereinafter).

In the practice of this invention, at least one cysteine residue is introduced in that portion of the IL-3, G-CSF or EPO polypeptide where modification via a sulfhydryl reactive compound is

desired. The cysteine residue or residues are so introduced by genetic engineering methods as described below. Novel cysteine residues may be engineered into the polypeptide for example, by simple insertion of a cysteine codon into the DNA molecule at the desired site or by converting a desirably located asparagine or other codon to a cysteine codon. Convenient methods for site specific mutagenesis or DNA synthesis for producing a DNA molecule encoding the desired CAV, expression in procaryotic or eucaryotic host cells of the DNA molecule so produced, and recovery of the CAV produced by such expression are also disclosed.

The CAVs of this invention retain useful biological properties of the natural or previously known protein and may thus be used for applications identified for the non-modified parent. Modification with such sulfhydryl reactive compounds, however, is preferred. Such biologically active, modified CAVs can be produced in homogeneous compositions which, it is contemplated, will provide improved pharmacokinetic profiles, immunogenicity profiles, and/or solubility characteristics relative to the parent polypeptides. Furthermore, CAVs may enable the formation of multimeric forms of the normally monomeric polypeptide with the same, albeit improved characteristics. Multimeric CAVs also enable the formation of "hetero-conjugates"--i.e., two or more distinct polypeptides joined via the sulfhydryl groups of the added cysteine residues, e.g., IL-3 joined to EPO or IL-3 joined to G-CSF.

Biological activity of the CAVs before or

after modification with the sulfhydryl reactive compounds may be determined by standard in vitro or in vivo assays conventional for measuring activity of the parent polypeptide. Alternatively, we provide herein a "small scale" screening method wherein successful Cys modification and attachment of the sulfhydryl reactive compound may be tested.

Selective and homogeneous modification of the CAVs with sulfhydryl reactive compounds is possible since such compounds will covalently bond primarily only to the cysteine residue(s) in the CAV. Secondary reactivity at His, Lys and Tyr residue(s) may be observed, depending on the choice of sulfhydryl reactive compound, but at a significantly lower rate. The modified CAVs so produced may then be recovered, and if desired, further purified and formulated into pharmaceutical compositions by conventional methods.

Sulfhydryl reactive compounds include compounds such as polyalkylene glycol, e.g. polyethylene and polypropylene glycol, as well as derivatives thereof, with or without coupling agents or derivatization with coupling or activating moieties, for example, with thiol, triflate, tresylate, aziridine or oxirane, or preferably with S-pyridyl or maleimide moieties. Compounds such as S-Pyridyl Monomethoxy PEG and Maleimido Monomethoxy PEG are exemplary. Additionally, sulfhydryl reactive compounds include, but are not limited to, charged or neutral polymers of the following types: dextran, colominic acids or other carbohydrate based polymers, polymers of amino acids and biotin derivatives, resulting in a protein modified with this well known affinity reagent often used for

antibody based assays.

Briefly, the method comprises reacting the CAV with a sulfhydryl reactive compound under suitable conditions, preferably non-denaturing conditions, and in sufficient amounts permitting the covalent attachment of the sulfhydryl reactive compound to the introduced cysteine residue(s) present in the polypeptide backbone of the CAV. The reaction may be reducible or non-reducible; and generally, the amount of sulfhydryl reactive compound used should be at least equimolar to the number of cysteines to be derivatized, although use of excess sulfhydryl reactive compound is preferred, both to improve the rate of reaction and to insure consistent modification at all reactive sites. The modified CAV produced, may then be recovered, purified and formulated by conventional methods. See e.g., WO 87/00056 and references cited therein.

Other aspects of the present invention include therapeutic methods of treatment and therapeutic compositions which employ the modified CAVs of the present invention, either alone or with other lymphokines, hematopoietins and/or growth factors, such as granulocyte macrophage colony-stimulating factor (GM-CSF), macrophage colony-stimulating factor (M-CSF), IL-1, IL-2, IL-4, IL-5, IL-6 and IL-10. These methods and compositions take advantage of the improved pharmacokinetic properties of these modified CAVs to provide treatments, e.g., such as employing lower dosages of polypeptide, less frequent administration, lower immunogenicity and more desirable distribution, required for the therapeutic indications for the natural polypeptide.

Other aspects and advantages of the present

invention will be apparent upon consideration of the following detailed description of the invention, including illustrative examples of the practice thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is the human IL-3 gene construct for E. coli expression, having the polypeptide sequence shown of natural (wild type) human IL-3, plus an initiator methionine, as expressed in E.coli, with the amino acids numbered from the N-terminus for reference to the muteins discussed herein.

Fig. 2 is the human G-CSF gene construct for E. coli expression, having the polypeptide sequence shown of natural (wild type) human G-CSF, plus an initiator methionine, as expressed in E. coli with the amino acids numbered from the N-terminus for reference to the muteins discussed herein.

Fig. 3 is a chemically synthesized human EPO gene construct for E. coli expression, having the polypeptide sequence of natural (wild type) human EPO, plus an initiator methionine, as expressed in E. coli with the amino acids numbered from N-terminus for reference to the muteins discussed herein.

DETAILED DESCRIPTION OF THE INVENTION

The present invention involves the selective modification of IL-3, G-CSF and EPO for pharmaceutical use, to both enhance their pharmacokinetic properties and provide homogeneous compositions for human therapeutic use. Although human IL-3, DNA and peptide sequences are preferred

as the starting point in this invention as it relates to IL-3, any primate IL-3 is susceptible to use in the method of the invention, given the significant homology between e.g., human and gibbon species of the protein and DNA. See Leary et al., Blood (1982) 70: 1343-1348. The method for selectively modifying IL-3, G-CSF and EPO involves selecting locations in the polypeptide sequence for the attachment of sulfhydryl reactive compounds. This step may be accomplished by altering the amino acid sequence of the polypeptide by inserting cysteine residues at selected sites or by converting selected endogenous residues into cysteine residues. For example, the codons AAA or AAG, which code for lysine, can be changed to the codon TGC or TGT, which code for cysteine.

CAVs in accordance with this invention also include allelic variations in the protein sequence, i.e. sequence variations due to natural variability from individual to individual, or with other amino acid substitutions or deletions which still retain desirable biological properties of the parent.

All CAVs of this invention may be prepared by expressing recombinant DNA sequences encoding the desired variant in host cells, e.g. procaryotic host cells such as E. coli, or eucaryotic host cells such as yeast or mammalian host cells, using methods and materials, e.g. vectors, as are known in the art. Host cells containing and capable of expressing the CAV-encoding DNA are thus encompassed by this invention. DNA sequences encoding the variants may be produced synthetically or by conventional site-directed mutagenesis of DNA sequences encoding the protein or polypeptide or analogs thereof. Figure 1 shows the human IL-3

gene construct inserted in plasmid pAL-hIL3-781 and expressed in the E. coli K12 strain designated GI586. This strain containing the plasmid was deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 USA on April 19, 1989 and given accession number 67932. Other DNA sequences for natural primate IL-3 have been cloned and the DNA sequences, including cDNA sequences, and specific peptide sequences for the same have been published, in PCT application number US87/01702, published as WO 88/00598 on January 28, 1988, and are therefore known in the art. These DNA sequences have been deposited with the American Type Culture Collection and given accession numbers ATCC 67154, 67326, 67319 and its replacement 68042, and 40246. DNA sequences for natural G-CSF and EPO have been cloned and the sequences and their corresponding peptide sequences published and are therefore known in the art.

DNA molecules encoding natural human IL-3s, G-CSFs and EPOs therefore may be obtained (i) by cloning in accordance with the published methods, (ii) from the deposited plasmids, or (iii) by synthesis, e.g. using overlapping synthetic oligonucleotides based on the published sequences which together span the desired coding region. Such methods are known in the art. See the foregoing PCT application published as WO 88/00598 and PCT application number US88/00402 published as WO88/06161.

As mentioned above, DNA sequences encoding individual CAVs of this invention may be produced synthetically or by conventional site-directed mutagenesis of a DNA sequence encoding the parental